Maximin-Projection Learning for Optimal Treatment Decision with Heterogeneous Individualized Treatment Effects

Chengchun Shi, Rui Song and Wenbin Lu

North Carolina State University, Raleigh, USA.

Bo Fu

Fudan University, Shanghai, People's Republic of China.

Summary. A saline feature of data from clinical trials and medical studies is inhomogeneity. Patients not only differ in baseline characteristics, but also the way they respond to treatment. Optimal individualized treatment regimes are developed to select effective treatments based on patient's heterogeneity. However, the optimal treatment regime might also vary for patients across different subgroups. In this paper, we mainly consider patients heterogeneity caused by groupwise individualized treatment effects assuming the same marginal treatment effects for all groups. We propose a new maximin-projection learning for estimating a single treatment decision rule that works reliably for a group of future patients from a possibly new subpopulation. Based on estimated optimal treatment regimes for all subgroups, the proposed maximin treatment regime is obtained by solving a quadratically constrained linear programming (QCLP) problem, which can be efficiently computed by interior-point methods. Consistency and asymptotic normality of the estimator is established. Numerical examples show the reliability of the proposed methodology.

Keywords: Heterogeneity; Maximin-projection learning; Optimal treatment regime; Quadratically constrained linear programming.

1. Introduction

Data from clinical trials and medical studies are often characterized by some degree of inhomogeneity. Patients not only differ in baseline characteristics, but also the way they respond to the treatment. There have been increasing interest in developing individualized optimal treatment regimes (OTRs) to account for patients' heterogeneity in response to treatment and to achieve the best treatment effect for individual patients. Some common methods for estimating OTRs include Q-learning (Watkins and Dayan, 1992; Chakraborty et al., 2010), A-learning (Robins et al., 2000; Murphy, 2003) and value search methods which directly search OTRs by maximizing the estimated value function

(Zhang et al., 2012; Zhao et al., 2012). However, the OTRs may also vary for patients from different subpopulations. This is typically the case in meta analysis, where we combine the results of multiple studies conducted at different locations or times. One motivating example is from a multi center randomised controlled trial as studied in Tarrier et al. (2004). The goal is to examine the effectiveness of cognitive-behavioural therapy for patients with early schizophrenia. Patients can be classified into three groups according to their treatment centres (Manchester, Liverpool and North Nottinghamshire). As we can see in Section D in the supplementary article, the group-wise OTRs can vary across different centres. Another example is from an observational study for investigating the influence of early disease modifying antirheumatic drug (DMARD) treatment on patients with recent onset inflammatory polyarthritis (Farragher et al., 2010). According to patients' enrollment time, they can be classified into three groups. As studied in Section 6, the group-wise OTRs can vary across different enrollment periods. The heterogeneity in OTRs may be explained by the differences in characteristics of treatment setting across subgroups. For instance, in the schizophrenia example, the strength of therapeutic alliance between therapist and patient, the adherence to treatment protocols and the quality of treatment provided can vary from one treatment centre to another (Dunn and Bentall, 2007); in the inflammatory polyarthritis example, there are more use of hydroxychloroquine for the methotrexate combination strategy in recruitment time group 3 (1997-2000) than in group 1 (1990-1992) or group 2 (1993-1996) as hydroxychloroquine was increasingly used in the UK before anti-tumour necrosis factor therapy was introduced to treat rheumatoid arthritis in 2001. Moreover, these characteristics are often unobserved or partially observed, and they may explain the interaction between subgroups and OTRs.

The aim of this paper is to propose a reliable OTR for new patients based on the observed data from different groups with heterogeneity in optimal treatment decision. The group of new patients may differ from any of the currently observed groups in terms of optimal treatment decision. For example, compared with existing data, the group of new patients, who come from a new treatment centre, may have a different OTR because of different strength of therapeutic alliance or different quality of treatment provided in the new treatment centre. Therefore, the true OTR for the group of new patients is not estimable at all based on the observed data, and any of the group-wise OTRs may not be the best choice. The challenge becomes how to derive a meaningful and reliable treatment regime that can take into account the heterogeneity in optimal treatment decision for different groups of patients. One simple approach is to pool the data of different groups together and obtain the "pooled" OTR based on the pooled data. Another method is to first obtain the OTR for each group, and then aggregate the group-wise OTRs in certain ways. Random effects meta-analysis (DerSimonian and Laird, 1986) is commonly used to combine subject-specific studies. Using its multivariate extensions

(cf. Jackson et al., 2010; Chen et al., 2012), we can aggregate the groupwise OTRs based on random effects models. The resulting OTR is similar to the "pooled" OTR when we have large numbers of subgroup patients. These OTRs maybe reasonable choices when the OTRs for different groups do not vary much. However, when there is certain degree of heterogeneity in OTRs across different groups as demonstrated in the toy example given in the next section, these OTRs are uniformly worse than the proposed OTR for any of the groups. One possible reason is that these OTRs for different groups may assign the same patient to different treatments and thus their effects are averaged out when pooling the data from different groups.

Bühlmann and Meinshausen (2016) and Meinshausen and Bühlmann (2015) considered a maximin criteria which has a nice characterization in linear models and proposed to use maximin aggregation (magging) to obtain the maximin estimator. Their proposed estimator is shown to be more robust than the pooled estimator in linear regression. The key idea of the maximin criteria is to find an estimator that works the best under the worst-case scenario. In optimal treatment decision, the percentage of making the correct decision (PCD) and value function are two commonly used measures to evaluate the effectiveness of a treatment regime. A natural maximin criteria for optimal treatment decision is to find an OTR that maximizes the minimum PCD or the minimum value function of all groups. Such a maximin OTR is appealing due to its nice interpretation and robustness. However, it is hard to implement in practice due to the following reasons. First, the PCD of a treatment regime is generally not estimable from data since the true OTR is unknown. Second, the empirical estimator of the value function as studied in Zhang et al. (2012) is non-smooth and non-concave, thus the estimation of the associated maximin OTR is not feasible.

In this paper, we propose a novel maximin-projection learning (MPL) to aggregate linear OTRs across different groups. Specifically, the proposed maximin-projection learning finds a linear decision rule that maximizes the minimum "inner product" between the vectors of regression parameters in the linear rule and the group-wise linear OTRs. We show that under certain model assumptions, the OTR obtained by the maximin-projection learning maximizes the minimum percentage of making the correct decision and value function of different groups, i.e. achieve the desired maximin properties. In addition, the corresponding estimation procedure can be represented as a linear programming problem with a quadratic constraint (Lee et al., 2016), which can be efficiently solved in $O(Gs^2 + s^3)$ flops. Here G denotes the number of groups and s the dimension of baseline covariates. Consistency and the asymptotic distribution of the corresponding maximin-projection estimators are established. Such kind of asymptotic results are rarely studied in the literature. To derive such asymptotic properties, we establish a necessary and sufficient condition for the existence and uniqueness of the population maximin-projection parameters and obtain a closed-form expression for the resulting estimator.

The rest of the paper is organized as follows. We introduce the model, notations and assumptions in Section 2. We also provide a heuristic comparison between the maximin OTR, the pooled OTR and the OTR based on random effects models with a toy example. In Section 3, we formally introduce the proposed maximin-projection learning including its statistical interpretation and geometrical characterization. Section 4 presents the estimating procedure of the maximin-projection estimator and the associated asymptotic properties. Simulation studies to evaluate the empirical performance of the proposed maximin OTR are conducted in Section 5. We apply our method to a real examples in Section 6, followed by a Conclusion Section. Proof of theorem 2 is provided in Section A. Other proofs and additional numerical studies are given in the supplementary article.

2. Preliminaries and a toy example

2.1. Preliminaries

For simplicity, we consider a single stage study with two treatments. Let Y denote a patient's response of interest, the larger the better by convention, $A \in \mathcal{A} = \{0, 1\}$ the treatment received by the patient and X the associated s-dimensional vector of baseline covariates. In addition, let $Y^*(0)$ and $Y^*(1)$ denote the potential outcomes that a patient would get if he or she was given treatment 0 and 1, respectively. A treatment regime d is a deterministic function that maps a patient's covariates to $\{0, 1\}$. Define the potential outcome

$$Y^*(d) = Y^*(1)d(X) + Y^*(0)\{1 - d(X)\},$$

representing the response that a patient would get if treated according to the regime d. The optimal treatment regime is defined as the regime d^{opt} that maximizes $\mathrm{E}\{Y^*(d)\}$. Under the stable unit treatment value assumption (SUTVA) and no unmeasured confounders assumption (Rubin, 1974), the optimal treatment regime can be written as $d^{opt}(x) = I\{C(x) > 0\}$ where

$$C(x) = E(Y|A = 1, X = x) - E(Y|A = 0, X = x).$$

Function $C(\cdot)$ is referred to as the contrast function. In practice, for simplicity, we may assume the contrast takes a linear form, i.e, $C(x) = \beta^T x + c$. To take population heterogeneity into account, we assume that the contrast function varies for patients from different groups. Specifically, we assume there are G groups of patients and consider the following semiparametric model:

$$Y_g = h_g(X_g) + A_g(\beta_g^T X_g + c_g) + e_g, \qquad g = 1, \ldots, G. \tag{1}$$

where $\mathrm{E}(e_g|X_g, A_g) = 0$. In Model (1), Y_g, A_g and $X_g \in \mathbb{R}^s$ stand for the response, the treatment and the covariates of patients in Group g, respectively, and h_g denotes the unspecified baseline function in Group g. Without loss of generality, we further assume all covariates X_g are standardized

to have zero mean and identity covariance matrix. Otherwise, we consider variable transformation $X_g^* = \Sigma_g^{-1/2}(X_g - \mu_g)$, $\beta_g^* = \Sigma^{1/2}\beta_g$, $c_g^* = c_g + \mu_g^T\beta_g$ where $\mu_g = \mathrm{E}(X_g)$ and $\Sigma_g = \mathrm{cov}(X_g)$. Then Model (1) can be represented as $Y_g = h_g^*(X_g^*) + A_g(\beta_g^{*T}X_g^* + c_g^*) + e_g$, for some function h_g^*. The parameter c_g stands for the marginal treatment effects (average causal effects) after adjusting covariates. Mathematically, we have

$$c_g = \mathrm{E}\{Y_g^*(1)\} - \mathrm{E}\{Y_g^*(0)\}.$$

When $c_g > 0$, treatment 1 is generally better for patients in Group g. The vector β_g describes individualized treatment effects. For patients in Group g with covariates x, the larger $\beta_g^T x$, the more benefits he or she receives if assigned to treatment 1.

Define $\pi_g(x) = \Pr(A_g = 1|X_g = x)$ as the propensity score in Group g. Model (1) allows h_g and π_g to vary across groups, which we refer to baseline effect heterogeneity and treatment assignment heterogeneity respectively. These sources of heterogeneity are not related to treatment decisions since they do not appear in the contrast function. The following sources of groupwise heterogeneity will affect decision making: the marginal treatment effects c_g and the individualized treatment effects β_g. In this paper, we mainly focus on heterogeneity caused by different β_g's. We assume $c_1 = \cdots = c_G = c_0$ for some c_0, that is, the same marginal treatment effect for all groups.

To introduce the pooled and the maximin optimal treatment regime, we need some optimality criterion. Here, we consider the difference of patient's mean response (value function) between a regime $d(x) = I(\beta^T x > -c)$ and $d_0(x) = 0$, which assigns all patients to treatment 0. Specifically, the difference of value functions is defined as

$$\mathrm{VD}_g(\beta, c) = \mathrm{E}\{Y_g^*(d)\} - \mathrm{E}\{Y_g^*(d_0)\} = \mathrm{E}\{(X_g^T\beta_g + c_0)I(X_g^T\beta > -c)\}.$$

In this section, for illustrative purposes only, we consider a special case with $c_0 = c = 0$. A general discussion will be given in the next section. When the distributions of X_gs are the same across groups, we can represent $\mathrm{VD}_g(\beta, 0)$ as

$$\mathrm{VD}(\beta, \beta_g) = \mathrm{E}\{(X_g^T\beta_g)I(X_g^T\beta > 0)\}.$$

We assume the same number of patients across all groups. Then, the pooled optimal treatment regime is defined as $d_P^{opt}(x) = I(x^T\beta^P > 0)$ where

$$\beta^P = \arg\max_{\|\beta\|_2=1} \frac{1}{G}\sum_{g=1}^{G} \mathrm{VD}(\beta, \beta_g), \tag{2}$$

and the maximin optimal treatment regime is defined as $d_M^{opt}(x) = I(x^T\beta^M > 0)$ where

$$\beta^M = \arg\max_{\|\beta\|_2=1} \min_{g\in\{1,\ldots,G\}} \mathrm{VD}(\beta, \beta_g). \tag{3}$$

We add the L_2 constraint on β to make β^P and β^M identifiable. Therefore, the pooled optimal treatment regime aims to maximize the average value difference while the maximin optimal treatment regime aims to maximize the minimum value difference in G groups, i.e. maximize the reward of the worst-case scenario.

The random effects meta-analyses assume the following model for β_g's:

$$\beta_g = \beta_0 + \varepsilon_g,$$

where ε_g's are independent and satisfy $\mathrm{E}(\varepsilon_g) = 0$, $\mathrm{cov}(\varepsilon_g) = \Omega_0$ for all g. For any subgroup estimators $\hat{\beta}_1, \ldots, \hat{\beta}_G$ with $\mathrm{cov}(\hat{\beta}_g) = \Omega_g$, the aggregated estimator is given by

$$\hat{\beta}^R = \left(\sum_{g=1}^{G} (\hat{\Omega}_g + \hat{\Omega}_0)^{-1} \right)^{-1} \left(\sum_{g=1}^{G} (\hat{\Omega}_g + \hat{\Omega}_0)^{-1} \hat{\beta}_g \right),$$

where $\hat{\Omega}_g$'s and $\hat{\Omega}_0$ denote some estimators for Ω_g's and Ω_0. Given sufficiently many observations, we have $\|\hat{\beta}_g - \beta_g\|_2 \xrightarrow{P} 0$ and $\|\hat{\Omega}_g\|_2 \xrightarrow{P} 0$. As a result, we have

$$\hat{\beta}^R \xrightarrow{P} \left(\sum_{g=1}^{G} (\hat{\Omega}_0)^{-1} \right)^{-1} \left(\sum_{g=1}^{G} (\hat{\Omega}_0)^{-1} \beta_g \right) = \frac{1}{G} \sum_{g} \beta_g \equiv \beta^R. \tag{4}$$

The corresponding optimal treatment regime is defined as $d_R^{opt}(x) = I(x^T \beta^R > 0)$.

More generally, we can treat the parameters β_g in the group-specific contrast function as a multivariate random variable and assume that the parameters β_g's of training groups are generated according to some distribution F_b, either continuous or discrete, and let H_b denote the support of F_b. Then, we define β^R, β^P and β^M as

$$\begin{aligned} \beta^R &= \mathrm{E}_{\mathrm{train},b}(b), \\ \beta^P &= \arg \max_{\|\beta\|_2 = 1} \mathrm{E}_{\mathrm{train},b}\{\mathrm{VD}(\beta, b)\}, \\ \beta^M &= \arg \max_{\|\beta\|_2 = 1} \min_{b \in H_b} \mathrm{VD}(\beta, b), \end{aligned}$$

where the expectation $\mathrm{E}_{\mathrm{train},b}$ is taken with respect to F_b. Definitions in (2), (3) and (4) correspond to the special case where F_b only takes values in $\{\beta_1, \ldots, \beta_G\}$ with an equal probability. Our objective is to minimize $\mathrm{E}_{\mathrm{test},b}\{\mathrm{VD}(\beta, b)\}$, where $\mathrm{E}_{\mathrm{test},b}$ is taken with respect to G_b, the distribution of β_g for future groups of patients.

2.2. A toy example

Recall that s is the dimension of X_g. For illustration, we take $s = 2$, and assume that patients' baseline covariates are generated independently from a standard normal distribution. Since $\|\beta\|_2 = 1$, after

some calculation, we have

$$\begin{aligned}\mathrm{VD}(\beta,\beta_g) &= \mathrm{E}\{X_g^T\beta_g I(X_g^T\beta>0)\} = \mathrm{E}\{(X_g^T\beta_g-\beta_g^T\beta X_g^T\beta+\beta_g^T\beta X_g^T\beta)I(X_g^T\beta>0)\}\\ &= \beta_g^T\beta\mathrm{E}\left\{X_g^T\beta I(X_g^T\beta>0)\right\} = \beta_g^T\beta\frac{1}{\sqrt{2\pi}}.\end{aligned}$$

The first equality in the second line is due to the independence between $X_g^T\beta_g-\beta_g^T\beta X_g^T\beta$ and $X_g^T\beta$. Hence, we obtain

$$\begin{aligned}\beta^P &= \arg\max_{\|\beta\|_2=1}\frac{1}{G}\sum_{g=1}^{G}\beta^T\beta_g = \frac{\sum_g\beta_g}{\|\sum_g\beta_g\|_2},\\ \beta^M &= \arg\max_{\|\beta\|_2=1}\min_{g\in\{1,\ldots,G\}}\beta^T\beta_g.\end{aligned} \tag{5}$$

Therefore, β^P is proportional to β^R which equals a simple average of all subgroup parameters, while β^M maximizes its minimum inner product across different β_g's. When all β_g's have the same L_2 norm, β^M becomes

$$\beta^M = \arg\max_{\|\beta\|_2=1}\min_{g\in\{1,\ldots,G\}}\frac{\beta^T\beta_g}{\|\beta_g\|_2}, \tag{6}$$

or equivalently

$$\beta^M = \arg\min_{\|\beta\|_2=1}\max_{g\in\{1,\ldots,G\}}\angle(\beta,\beta_g), \tag{7}$$

where $\angle(a,b)=\arccos(a^Tb)$ stands for the angle between two vectors. The equivalence between (6) and (7) is due to the monotonicity of the arccos function. In (6) or (7), β^M is defined to maximize (minimize) the minimum correlation (maximum angle) between all subgroup coefficients. Such formulation is referred to as the maximin correlation approach in the classification literature (c.f, Avi-Itzhak et al., 1995; Lee et al., 2016). In general, we weight the correlation $\beta^T\beta_g/\|\beta_g\|_2$ by the L_2 norm of β_g. The β^M defined in (5) is more informative since it not only takes the heterogeneity due to different directions $\beta_g/\|\beta_g\|_2$ into consideration, but different magnitudes $\|\beta_g\|_2$ as well.

Since β^P is proportional to β^R, the VD under d_P^{opt} is the same as d_R^{opt}. Therefore, in the following, we focus on comparing β^P with β^M. We set $G=4$ and assume $\|\beta_g\|_2=1$, $g=1,2,3,4$. Since $s=2$, we represent each β_g as $\beta_g=\{\cos(\psi_g),\sin(\psi_g)\}$ with $\psi_g\in[0,\pi)$. The parameter ψ_g is the angle between β_g and the x-axis in a 2-dimensional coordinate system. In this special case, β^M lies on the bisector of the largest angles formed by all β_g's and it can be shown that $\beta^M=\{\cos(\psi^M),\sin(\psi^M)\}$ where

$$\psi^M = \frac{1}{2}\left(\psi_{(1)}+\psi_{(4)}\right),$$

$\psi_{(1)}$ and $\psi_{(4)}$ denote the smallest and largest angles of ψ_g's. Similarly define $\beta^P=\{\cos(\psi^P),\sin(\psi^P)\}$. We set $\psi_1=0°$, $\psi_2=15°$, $\psi_3=70°$ and $\psi_4=90°$. Consider the following leave-one-group-out cross

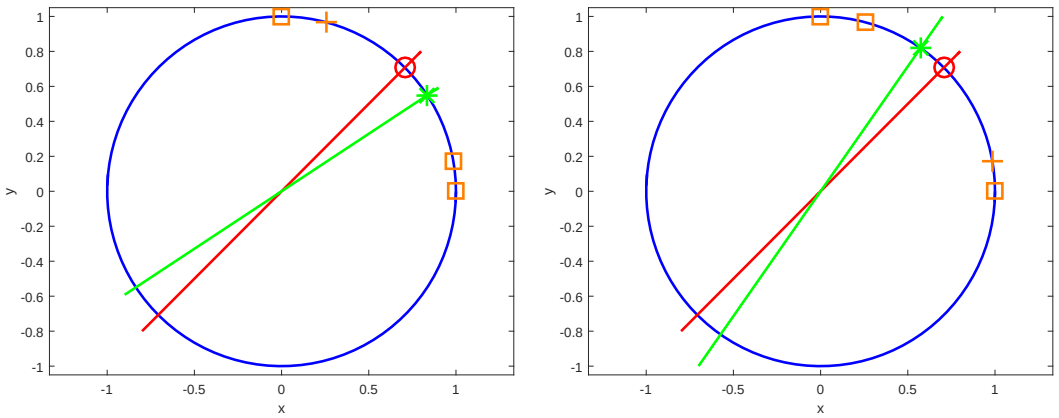
validation procedure. For the ith round, we choose the ith group as the testing group, and obtain β^P and β^M based on the remaining 3 groups. Then we evaluate the value difference of the pooled and maximin OTRs based on the ith group. In other words, we set F_b to be a discrete distribution that takes value on $\{\beta_1, \ldots, \beta_4\}/\{\beta_i\}$ with equal probability, and G_b a degenerate distribution that concentrates on β_i. Table 1 summarizes the results.

Table 1: Different combinations of training groups and the corresponding ψ^P, ψ^M, and their value differences on the testing group

Training groups	ψ^M (deg)	ψ^P (deg)	ψ_{test} (deg)	VD($\beta^M, \beta_{\text{test}}$)	VD($\beta^P, \beta_{\text{test}}$)
$(1,2,3)$	35	27.44	90	0.23	0.18
$(1,2,4)$	45	32.63	70	0.36	0.32
$(1,3,4)$	45	55.32	15	0.35	0.30
$(2,3,4)$	52.5	59.25	0	0.24	0.20

From Table 1, we can see that for all four cases, the value differences of the maximin optimal treatment regime are uniformly larger than those of the pooled optimal treatment regime on the testing groups. To illustrate the idea graphically, we plot β^P (denoted by the snow symbol), β^M (denoted by the circle symbol), and β_g of the training (denoted by the square symbol) and testing (denoted by the plus symbol) groups for the second and third cases in Figure 1, where the left panel is for the second case and the right one is for the third case. For both cases, β^M is closer to β_g of the testing groups, while β^P is pulled towards the area where most β_g's of the training groups locate due to the averaging effect.

Fig. 1: Plots of β^P (denoted by the snow symbol), β^M (denoted by the circle symbol), and β_g of the training (denoted by the square symbol) and testing groups (denoted by the plus symbol) for the second (left panel) and third (right panel) cases.



3. Maximin-projection learning

We now formally introduce our maximin projection treatment regime. Based on model (1) and the common marginal treatment effect assumption, the optimal treatment regime for the gth subgroup is $d_g^{opt}(x) = I(x^T\beta_g > -c_0)$. Here, our goal is to find a single treatment regime $d_M^{opt}(x) = I(x^T\beta^M > -c^M)$ with $\|\beta^M\|_2 = 1$ that performs uniformly well for heterogeneous data. Motivated by the toy example in the previous section, our proposed maximin-projection learning is aim to find

$$\beta^M = \arg \max_{\beta:\|\beta\|_2=1} \min_{g\in\{1,\ldots,G\}} \beta^T\beta_g.$$

3.1. Statistical interpretation

In this subsection, we show that the maximin projection, represented by β^M, has two nice statistical interpretations in terms of maximizing the minimum PCD and value difference (VD). Specifically, in group g, the PCD of a treatment regime $d(x) = I(x^T\beta > -c)$ is defined as

$$\mathrm{PCD}_g(\beta, c) = 1 - \mathrm{E}\left\{|I(X_g^T\beta > -c) - I(X_g^T\beta_g > -c_0)|\right\},$$

and the VD is defined as

$$\mathrm{VD}_g(\beta, c) = \mathrm{E}[Y_g^*\{I(X_g^T\beta > -c)\}] - \mathrm{E}\{Y_g^*(0)\} = \mathrm{E}\left\{(X_g^T\beta_g + c_0)I(X_g^T\beta > -c)\right\}.$$

Here, the larger PCD and VD values, the better the treatment regime $d(x)$ approximates the groupwise optimal treatment regime $d_g^{opt}(x)$.

Based on the defined PCD and VD, for any fixed constant c, we consider the following maximin treatment regimes: $d_1(x) = I(x^T\beta_{(1)}^M > -c)$ where

$$\beta_{(1)}^M = \arg \max_{\beta:\|\beta\|_2=1} \min_{g\in\{1,\ldots,G\}} \mathrm{PCD}_g(\beta, c), \tag{8}$$

and $d_2(x) = I(x^T\beta_{(2)}^M > -c)$ where

$$\beta_{(2)}^M = \arg \max_{\beta:\|\beta\|_2=1} \min_{g\in\{1,\ldots,G\}} \mathrm{VD}_g(\beta, c). \tag{9}$$

Remark 3.1. *The two maximin treatment regimes, defined by $\beta_{(1)}^M$ and $\beta_{(2)}^M$, are appealing for their nice statistical interpretations. However, we note that the definition of $\beta_{(1)}^M$ involves unknown parameters. The empirical estimators of VD are of non-smooth and non-concave functional forms of the corresponding estimators. Therefore, their estimations are not feasible and they may not be practically useful.*

Remark 3.2. *It is worth noting that $\beta_{(1)}^M$ would be meaningless when not all $\|\beta_g\|_2$'s are the same. This is because PCD only measures the similarity between the overall and groupwise optimal treatment*

decisions, but does not account for the magnitude of groupwise contrast function. When $\|\beta_g\|_2$'s are not the same, the L_2 norm of groupwise contrast function $\{E(X_g^T\beta_g + c_0)^2\}^{1/2}$ would be different. This implies that PCDs are not comparable across different groups. In comparison, VD is a better criterion since it takes both the sign and magnitude of contrast function into consideration. Below, under some conditions, we establish the equivalence between these two maximin treatment regimes and our proposed maximin-projection treatment regime.

THEOREM 1 (EQUIVALENCE OF β^M AND $\beta^M_{(1)}$). *Assume that X_g's are i.i.d. spherically distributed, and all $\|\beta_g\|_2$'s are the same. Then, for any fixed c,*

$$\beta^M = \arg\max_{\|\beta\|_2=1} \min_{g\in\{1,\ldots,G\}} PCD_g(\beta, c).$$

THEOREM 2 (EQUIVALENCE OF β^M AND $\beta^M_{(2)}$). *Assume X_g's are i.i.d. spherically distributed. Then, for any fixed c,*

$$\beta^M = \arg\max_{\|\beta\|_2=1} \min_{g\in\{1,\ldots,G\}} VD_g(\beta, c).$$

REMARK 3.3. *Theorems 1 and 2 require X_g to have a spherical distribution (see Definition F.1), which is a rich class of symmetric multivariate distributions (see Fang et al., 1990).*

The definition of β^M has nice statistical interpretations. However, it has two drawbacks. First, when $F_0 \equiv \max_{\|\beta\|_2=1} \min_g \beta^T\beta_g < 0$, the uniqueness of β^M is not guaranteed. This may cause identifiability issues when we establish properties of the corresponding estimators. In addition, the optimization problem in (5) is not concave. This can make the implementation of the estimating procedure infeasible.

To address these concerns, we define

$$\beta^M_{(0)} = \arg\max_{\|\beta\|_2\le 1} \min_{g\in\{1,\ldots,G\}} \beta^T\beta_g. \tag{10}$$

Compared to β^M, it replaces the feasible set $\|\beta\|_2 = 1$ with a closed convex set $\|\beta\|_2 \le 1$. Lemma 1 below states that $\beta^M_{(0)}$ is well defined, when $F_0 \neq 0$. Moreover, the optimization problem (10) is concave, which can be easily implemented.

LEMMA 1. *The maximin-projection estimator $\beta^M_{(0)}$ always exists. Moreover, when $F_0 \neq 0$, $\beta^M_{(0)}$ is unique.*

REMARK 3.4. *The existence of $\beta^M_{(0)}$ is guaranteed by the continuity of the objective function $F(\beta) = \min_{g\in\{1,\ldots,G\}} \beta^T\beta_g$, boundedness and closeness of the feasible set $\beta : \|\beta\|_2 \le 1$. Its uniqueness is a*

byproduct of lemma 3, which is stated in the next subsection. When $F_0 = 0$, $\beta^M_{(0)}$ *is not unique and the set of solutions is given by*

$$\{a\beta : a \in [0,1], \|\beta\|_2 = 1, \max_{\|\beta\|_2=1} \min_{g\in\{1,\ldots,G\}} \beta^T\beta_g = 0\}.$$

The problem of estimating $\beta^M_{(0)}$ *then becomes non-regular and all the large sample theories about the maximin estimator fail (see Section 4).*

Define $G_0 = \max_{\|\beta\|_2\le 1} \min_g \beta^T\beta_g$. It is obvious that $G_0 \ge 0$. In addition, $G_0 > 0$ if and only if $F_0 > 0$. When $G_0 = 0$, we can set $\beta^M_{(0)} = 0$, which leads to a trivial regime by assigning the same treatment to all patients. From now on, we focus on the situation when $G_0 > 0$. In this case, we have $\beta^M = \beta^M_{(0)}$. Define

$$c^M_{(0)} = c_0/G_0.$$

Note that $c^M_{(0)}$ and c_0 are sign equivalent. Our maximin-projection OTR is given by

$$d^{opt}_M(x) = I(x^T\beta^M_{(0)} > -c^M_{(0)}).$$

THEOREM 3. *Under conditions of theorem 1, if* $G_0 > 0$, *we have*

$$c^M_{(0)} = \arg\max_c \min_{g\in\{1,\ldots,G\}} PCD_g(\beta^M_{(0)}, c).$$

THEOREM 4. *Under conditions of theorem 2, if* $G_0 > 0$, *we have*

$$c^M_{(0)} = \arg\max_c \min_{g\in\{1,\ldots,G\}} VD_g(\beta^M_{(0)}, c).$$

Together with theorems 1 and 2, theorems 3 and 4 suggest that the treatment regime $d^{opt}_M(x)$ maximizes the minimum PCD and the minimum VD among different groups.

3.2. Geometrical characterization

In this subsection we give a geometrical view of $\beta^M_{(0)}$ when $G_0 > 0$. Findings in this subsection are similar in rationale with the results in Avi-Itzhak et al. (1995). However, we generalize their results by getting rid of the unit L_2-norm condition $\|\beta_g\|_2 = 1$ and allowing the set of vectors $\{\beta_1, \ldots, \beta_G\}$ to be linear dependent, which is the case when $s \ge G$.

We first introduce some notation. For an arbitrary $s \times G$ matrix Ψ and a set $K \subseteq \{1, \ldots, G\}$, let Ψ_K denote the submatrice of Ψ formed by columns in K. Define the equicorrelated points set

$$\mathrm{E}_K(\Psi) = \left\{t \in \mathbb{R}^s | t^T\Psi_j = t^T\Psi_i, \forall i, j \in K\right\},$$

and the optimal equicorrelated point

$$\mathrm{E}^\star_K(\Psi) = \arg\max_{\substack{t\in E_K(\Psi)\\ \|t\|_2=1}} \left\{t^T\Psi_i, \forall i \in K\right\},$$

where Ψ_i refers to the ith column vector of matrix Ψ. When $|K| = 1$ and $\Psi_K = \psi$, $\mathrm{E}^\star_K(\Psi) = \psi/\|\psi\|_2$. Readers can refer to Section B of the supplementary article for a detailed discussion on the equicorrelated points set and the optimal equicorrelated point.

For any matrix Ω, Let Ω^+ denote the Moore-Penrose matrix inverse of Ω and $C(\Omega)$ the column space of Ω. Let e denote a vector of ones. We have the following result.

LEMMA 2. *For any Ψ and $K \subseteq [1, \ldots, n]$, when $e \in C(\Psi_K^T)$, the optimal equicorrelated point of Ψ_K exists and is unique. Moreover, it takes the form*

$$E^\star_K(\Psi) = [e^T(\Psi_K^T\Psi_K)^+e]^{-1/2}\Psi_K(\Psi_K^T\Psi_K)^+e. \tag{11}$$

Define matrix $B = (\beta_1, \beta_2, \ldots, \beta_G)$ whose gth column is the subgroup parameter β_g.

LEMMA 3. *Assume $G_0 > 0$. Then there exists a unique nonempty set $K_0 \subseteq [1, \ldots, G]$ such that $\beta^M_{(0)} = E^\star_{K_0}(B)$ and $\min_{g\in K_0^c} {\beta^M_{(0)}}^T\beta_g > G_0$, where $K_0^c = [1, \ldots, G] - K_0$. Moreover, if the set of vectors $\beta_g, g \in K_0$ are linearly independent, then a necessary and sufficient condition for $\beta^M_{(0)} = E^\star_{K_0}(B)$ is that each element in the vector $(B^T_{K_0}B_{K_0})^{-1}e$ is nonnegative.*

We denote K_0 as the maximin optimal equicorrelated points set when $G_0 > 0$. In lemma 2, the condition $e \in C(\Psi_K^T)$ automatically holds when Ψ_K^T has full row rank. In lemma 3, we assume the set of vectors $\beta_g, g \in K_0$ are linearly independent. This implies the matrix $B^T_{K_0}$ has full row rank. As a result, we have $e \in C(B^T_{K_0})$.

In lemma 3, the non-negativity of $(B^T_{K_0}B_{K_0})^{-1}e$ is sufficient and necessary for $\beta^M_{(0)} = \mathrm{E}_{K_0}(B)$. Together with lemma 2, lemma 3 implies that $\beta^M_{(0)}$ is uniquely defined by

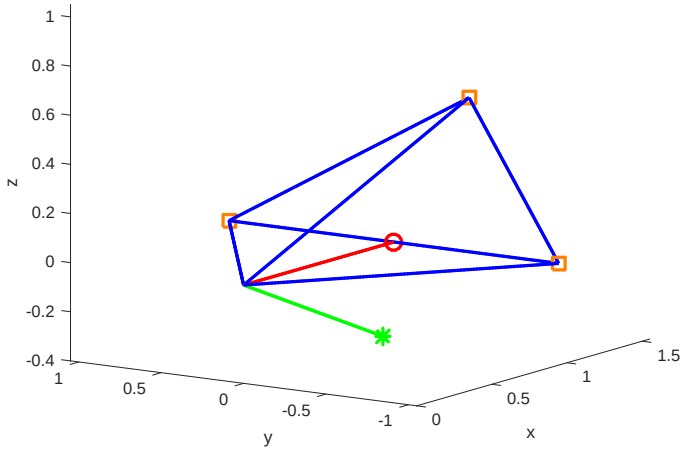
$$\beta^M_{(0)} = \mathrm{E}^\star_{K_0}(B) = [e^T(B^T_{K_0}B_{K_0})^{-1}e]^{-1/2}B_{K_0}(B^T_{K_0}B_{K_0})^{-1}e.$$

This implies $\mathrm{E}^\star_{K_0}(B)$ is proportional to $B_{K_0}(B^T_{K_0}B_{K_0})^{-1}e$ and can be represented as a linear combination of the column vectors in B_{K_0}. Geometrically, the non-negativity of $(B^T_{K_0}B_{K_0})^{-1}e$ requires $\mathrm{E}^\star_{K_0}(B)$ to lie in the convex cone of $\beta_g, g \in K_0$, i.e, $\{\sum_{g\in K_0} a_g\beta_g : a_g \geq 0, \forall g \in K_0\}$. To better understand lemma 3, in Figure 2, we take $s = 3$, $G = 3$ and $B = (\beta_1, \beta_2, \beta_3)$ where $\beta_1 = (1, 1, 0)$, $\beta_2 = (1, -1, 0)$ and $\beta_3 = (1.2, 0, 0.5)$. Both $\mathrm{E}^\star_{\{1,2,3\}}(B)$ and $\mathrm{E}^\star_{\{1,2\}}(B)$ satisfy the necessary conditions of lemma 3. While $\mathrm{E}^\star_{\{1,2\}}(B)$ lies in the convex cone of β_1 and β_2, $\mathrm{E}^\star_{\{1,2,3\}}(B)$ appears outside the convex cone of β_1, β_2 and β_3. Therefore, $\mathrm{E}^\star_{\{1,2\}}(B)$ satisfies the sufficient conditions of lemma 3 and $\mathrm{E}^\star_{\{1,2,3\}}(B)$ doesn't. As a result, we have $\beta^M_{(0)} = \mathrm{E}^\star_{\{1,2\}}(B)$.

4. Estimation procedure

The data are summarized as (Y_{gj}, A_{gj}, X_{gj}), for $g = 1, \ldots, G$, $j = 1, \ldots, m_g$, where m_g is the number of patients in Group g. We assume that the data are independent across $g = 1, \ldots, G$ and $j = 1, \ldots, m_g$.

Fig. 2: Plots of β_g (denoted by the square symbol), $\mathrm{E}^{\star}_{\{1,2,3\}}(B)$ (denoted by the snow symbol) and $\mathrm{E}^{\star}_{\{1,2\}}(B)$ (denoted by the circle symbol)



Based on the data, parameters $\beta_1, \ldots, \beta_G$ and c_0 in model (1) can be estimated with existing methods. In this paper, we implement with the popular Q-learning and A-learning and give a brief discussion on estimating these parameters in Section 4.2. Let $\hat{\beta}_1, \ldots, \hat{\beta}_G$ and $\hat{c}_0$ be the corresponding estimators. We propose to estimate $\beta^M_{(0)}$ by solving the following optimization problem:

$$\hat{\beta}^M = \arg \max_{\beta:\|\beta\|_2 \le 1} \min_{g \in \{1,\ldots,G\}} \beta^T \hat{\beta}_g. \tag{12}$$

Note that the objective function $\min_g \beta^T \hat{\beta}_g$ is concave in β and the region $\|\beta\|_2 \le 1$ is convex. Therefore, (12) is a tractable convex optimization problem. It can be further casted as a quadratic constraint linear programming (QCLP) problem, specifically, $\hat{\beta}^M$ is equivalent to the solution of

$$\begin{aligned} &\text{maximize} && t \in \mathbb{R} \\ &\text{subject to} && \beta^T \hat{\beta}_g \ge t, g = 1, \ldots, G \\ & && \beta^T \beta \le 1. \end{aligned}$$

The above optimization problem can be efficiently computed using existing softwares. Define $\hat{c}^M = \hat{c}_0 / \hat{G}_0$, where $\hat{G}_0 = \min_g \hat{\beta}_g^T \hat{\beta}^M$.

Given a group of future patients, denoted by $\{X_{G+1,j}\}_{j=1}^n$ their baseline covariates. We calculate $\hat{\mu}_{G+1} = \sum_{j=1}^n X_{G+1,j}/n$ and $\hat{\Sigma}_{G+1} = \sum_{j=1}^n (X_{G+1,j} - \hat{\mu}_{G+1})(X_{G+1,j} - \hat{\mu}_{G+1})^T/(n-1)$. The recommend treatment for the jth patient is given by

$$I\{(X_{G+1,j} - \hat{\mu}_{G+1})^T \hat{\Sigma}_{G+1}^{-1/2} \hat{\beta}^M > -\hat{c}^M\}.$$

4.1. Statistical properties

In this subsection we investigate the asymptotic properties of the maximin-projection estimator $\hat{\beta}^M$ obtained by solving the optimization problem (12). We first study the consistency of the estimator by assuming the following two conditions.

(C1.) Assume that $\hat{\beta}_1, \ldots, \hat{\beta}_G$ and $\hat{c}_0$ converge in probability to $\beta_1, \ldots, \beta_G$ and c_0, respectively.

(C2.) Assume that $F_0 \neq 0$. When $F_0 > 0$, assume that the column vectors in B_{K_0} are linearly independent and all elements in the vector $(B_{K_0}^T B_{K_0})^{-1} e$ are nonzero, where K_0 is the maximin optimal equicorrelated points set as defined previously.

REMARK 4.1. *Condition (C1) requires each subgroup estimator to be consistent. The condition $F_0 \neq 0$ in (C2) ensures the existence and uniqueness of $\beta_{(0)}^M$. Apparently, $\beta_{(0)}^M$ is not stable when F_0 approaches to 0, since its L_2 norm will change from 1 to 0. To ensure the stability of $\beta_{(0)}^M$ in the sense that it will not deviate too much when there are minor changes in the set of vectors $\beta_1, \ldots, \beta_G$, we would expect*

$$\|[\tilde{B}_{K_0}^T \tilde{B}_{K_0}]^+ - [B_{K_0}^T B_{K_0}]^+\|_2 \to 0, \tag{13}$$

as $\tilde{B}_{K_0} \to B_{K_0}$, where $\tilde{B} = (\tilde{\beta}_1, \ldots, \tilde{\beta}_G)$ represents the coefficient matrix with some disturbance. A sufficient condition to establish (13) *is that B_{K_0} is of full column rank, as assumed in Condition (C2). Lemma 2 suggests $\beta_{(0)}^M$ can be represented as $\omega_0^T B_{K_0}$, for some weight vector ω_0 proportional to $(B_{K_0}^T B_{K_0})^{-1} e$. Condition (C2) further assumes the weights are nonzero. Such a condition guarantees that for any coefficient matrix $\tilde{B} \to B$, K_0 is the optimal equicorrelated points set of $\tilde{B}$ as well.*

THEOREM 5 (CONSISTENCY). *Define $\hat{B} = (\hat{\beta}_1, \ldots, \hat{\beta}_G)$. Assume Conditions C1 and C2 are satisfied. Then with probability tending to 1, the estimator $\hat{\beta}^M$ is equal to*

$$\begin{cases} \{e^T (\hat{B}_{K_0}^T \hat{B}_{K_0})^{-1} e\}^{-1/2} \hat{B}_{K_0} (\hat{B}_{K_0}^T \hat{B}_{K_0})^{-1} e & \text{if } F_0 > 0, \\ 0 & \text{if } F_0 < 0. \end{cases}$$

In addition, assume there exist some $r_n^{(1)}, r_n^{(2)} \to 0$ such that $\max_{g \in K_0} \|\hat{\beta}_g - \beta_g\|_2 = O_p(r_n^{(1)})$ and $\hat{c}_0 = c_0 + O_p(r_n^{(2)})$. When $F_0 > 0$, we have $\|\hat{\beta}^M - \beta_{(0)}^M\|_2 = O_p(r_n^{(1)})$, $\hat{c}^M = c_{(0)}^M + O_p(r_n^{(1)} + r_n^{(2)})$.

REMARK 4.2. *Theorem 5 implies that $(\hat{\beta}^M, \hat{c}^M)$ is consistent as long as each subgroup estimator is consistent. The first part of the theorem follows as a consequence of lemma 3.*

Next, we study the asymptotic normality of the estimator. For notational simplicity, we assume $m_1 = \cdots = m_G = m$ and posit the following condition.

(C3.) Assume that for all $g \in K_0$, $\sqrt{m}(\hat{\beta}_g - \beta_g)$ and $\sqrt{m}(\hat{c}_0 - c_0)$ are jointly asymptotically normal with mean zero.

THEOREM 6 (ASYMPTOTIC NORMALITY). *Assume that Conditions C1–C3 hold, and that* $F_0 > 0$. *We have that* $\sqrt{m}(\hat{\beta}^M - \beta^M_{(0)})$ *and* $\sqrt{m}(\hat{c}^M - c^M_{(0)})$ *are jointly asymptotically normal with mean zero and some covariance matrix* V^M. *The expression of* V^M *is given in Appendix C.*

Since the expression of the asymptotic covariance matrix V^M is quite complicated, we propose to estimate it using a bootstrap method. Here, the bootstrap sampling is done within each subgroup. Specifically, we independently generate B bootstrap samples for each group $g = 1, \ldots, G$,

$$\left\{(Y^{(j)}_{g1}, A^{(j)}_{g1}, X^{(j)}_{g1}), \ldots, (Y^{(j)}_{gm}, A^{(j)}_{gm}, X^{(j)}_{gm})\right\},$$

$j = 1, \ldots, B$. For each j, we obtain estimators $\hat{\beta}^{(j)}$ and $\hat{c}^{(j)}$ based on the data

$$\left\{(Y^{(j)}_{11}, A^{(j)}_{11}, X^{(j)}_{11}), \ldots, (Y^{(j)}_{1m}, A^{(j)}_{1m}, X^{(j)}_{1m})\right\}, \ldots, \left\{(Y^{(j)}_{G1}, A^{(j)}_{G1}, X^{(j)}_{G1}), \ldots, (Y^{(j)}_{Gm}, A^{(j)}_{Gm}, X^{(j)}_{Gm})\right\}.$$

Confidence intervals of $\hat{\beta}^M$ and $\hat{c}^M$ are calculated based on quantiles of $(\hat{\beta}^{(1)}, \ldots, \hat{\beta}^{(B)})$ and $(\hat{c}^{(1)}, \ldots, \hat{c}^{(M)})$.

4.2. Estimation of group-specific regimes

In this subsection we discuss two popular approaches to obtain subgroup estimators $\hat{\beta}_g$ and $\hat{c}_0$.

EXAMPLE 1 (Q-LEARNING). *We estimate* β_g *and* c_0 *by modeling the Q-functions, which represent the conditional mean of the response given the covariates and the treatment. Specifically, the baseline function is assumed to have some parametric form* $h_g(x, \eta_g)$ *with parameter* η_g. *Then,*

$$Q_g(X_g, A_g; \beta_g, c_0, \eta_g) \equiv E(Y_g | A_g, X_g) = h_g(X_g, \eta_g) + A_g(X_g^T \beta_g + c_0), \quad g = 1, \ldots, G.$$

Since c_0 *is common across all subgroups, we propose to estimate* $\beta_1, \ldots, \beta_G$ *and* c_0 *by jointly solving the following set of estimating equations:*

$$\begin{aligned}
&\sum_j \frac{\partial h_g(X_{gj}, \eta_g)}{\partial \eta_g} \{Y_{gj} - Q_g(X_{gj}, A_{gj}; \beta_g, c_0, \theta_g)\} = 0, \quad g = 1, \ldots, G, \\
&\sum_j A_{gj} X_{gj} \{Y_{gj} - Q_g(X_{gj}, A_{gj}; \beta_g, c_0, \theta_g)\} = 0, \quad g = 1, \ldots, G, \\
&\sum_g \sum_j A_{gj} \{Y_{gj} - Q_g(X_{gj}, A_{gj}; \beta_g, c_0, \theta_g)\} = 0.
\end{aligned}$$

When the parametric models $h_g(x, \eta_g)$*'s are correctly specified, the resulting estimators* $\hat{\beta}_g$*'s and* $\hat{c}_0$ *are consistent and jointly asymptotically normal.*

EXAMPLE 2 (A-LEARNING). *Here, we posit some parametric model* $\pi_g(X, \alpha_g)$ *for the propensity score and* $h_g(X, \eta_g)$ *for the baseline function. The parameters* α_g*'s,* η_g*'s,* β_g*'s and* c_0 *are estimated*

by solving the following set of estimating equations:

$$\sum_j \frac{1}{\pi_g(X,\alpha_g)\{1-\pi_g(X,\alpha_g)\}}\frac{\partial \pi_g(X,\alpha_g)}{\partial \alpha_g}\{A_{gj}-\pi_g(X,\alpha_g)\}=0, \quad g=1,\ldots,G,$$

$$\sum_j \frac{\partial h_g(X_{gj},\eta_g)}{\partial \eta_g}\{Y_{gj}-h_g(X_{gj},\eta_g)-A_{gj}(X_{gj}^T\beta_g+c_0)\}=0, \quad g=1,\ldots,G,$$

$$\sum_j X_{gj}\{A_{gj}-\pi_g(X_{gj},\alpha_g)\}\{Y_{gj}-h_g(X_{gj},\eta_g)-A_{gj}(X_{gj}^T\beta_g+c_0)\}=0, \quad g=1,\ldots,G,$$

$$\sum_g\sum_j \{A_{gj}-\pi_g(X_{gj},\alpha_g)\}\{Y_{gj}-h_g(X_{gj},\eta_g)-A_{gj}(X_{gj}^T\beta_g+c_0)\}=0.$$

It can be shown that when either the propensity score or the baseline function for each group is correctly specified, the resulting estimators $\hat{\beta}_g$'s and $\hat{c}_0$ are consistent and jointly asymptotically normal. This is the so-called doubly robust property of the A-learning estimation.

5. Simulation studies

We consider four groups of patients. In each group, we generate 200 samples according to the following model

$$Y_{gj} = h(X_{gj}) + A_{gj}X_{gj}^T\beta_g + \varepsilon_{gj},$$

where $X_{gj} = (X_{gj}^{(1)}, X_{gj}^{(2)})^T \overset{iid}{\sim} N(0, I_2)$ and $\varepsilon_{gj} \overset{iid}{\sim} N(0, 0.25)$. Two baseline models are considered for h, including a linear model $h(X_{gj}) = 1 + 0.5X_{gj}^{(1)} + 0.5X_{gj}^{(2)}$ and a nonlinear model $h(X_{gj}) = 1 + \sin(0.5\pi X_{gj}^{(1)} + 0.5\pi X_{gj}^{(2)})$. We generate treatments from two propensity score models, a constant model, $\Pr(A_{gj} = 1) = 0.5$ and a probit model, $\Pr(A_{gj} = 1|X_{gj}) = \Phi(X_{gj}^{(1)} - X_{gj}^{(2)})$, where $\Phi(\cdot)$ is the standard normal cumulative distribution function. This yields four simulation settings.

We further consider two scenarios for the subgroup parameters to exhibit different degrees of heterogeneity. In the first scenario, we set $\beta_1^T = (2, 0), \beta_2^T = (2\cos(15°), 2\sin(15°)), \beta_3^T = (2\cos(70°), 2\sin(70°)), \beta_4^T = (0, 2)$. Hence, all β_g's have the same L_2 norm and their directions $\beta_g/\|\beta_g\|_2$ differ. For the second scenario, we choose subgroup parameters to have similar directions but allow their L_2 norms to vary. Specifically, $\beta_1^T = (2.2\cos(30°), 2.2\sin(30°))$, $\beta_2^T = (1.5\cos(45°), 1.5\sin(45°))$, $\beta_3^T = (2.2\cos(54°), 2.2\sin(54°))$, $\beta_4^T = (2\cos(60°), 2\sin(60°))$. It can be shown that $\beta_{(0)}^M = (\cos(45°), \sin(45°))$ and $c_{(0)}^M = 0$ for all scenarios.

We first obtain the subgroup estimators of β_g and c_0 using the A-learning estimating equations discussed in Section 4.2. Here, a logistic regression model is fitted for the propensity score and a linear model for the baseline function. As a result, both the propensity score model and the baseline model are correctly specified in the first setting; either of them is misspecified in the second and the third setting; while both are misspecified in the last setting. We then obtain the estimators $\hat{\beta}^M$ and

$\hat{c}^M$ using the proposed maximin-projection learning. Confidence intervals for the resulting estimators are obtained based on 600 bootstrap samples.

Table 2: Biases, standard deviations (in parenthesis) of $\hat{\beta}^M$, $\hat{c}^M$ and coverage probabilities (CP) of 95% Wald-type confidence intervals for $\beta^M_{(0)}$ and $c^M_{(0)}$.

Scenario 1	$\hat{\beta}^M_1$	$\hat{\beta}^M_2$	$\hat{c}^M$	CP for $\hat{\beta}^M_1$	CP for $\hat{\beta}^M_2$	CP for $\hat{c}^M$
Setting 1	−0.002(0.027)	0.001(0.027)	0.0003(0.024)	96.0%	96.0%	95.3%
Setting 2	−0.003(0.053)	−0.001(0.052)	0.001(0.045)	94.7%	94.7%	93.8%
Setting 3	−0.003(0.036)	0.001(0.035)	−0.0005(0.035)	96.2%	96.2%	94.5%
Setting 4	−0.003(0.068)	−0.004(0.068)	0.002(0.068)	96.0%	96.0%	95.0%
Scenario 2	$\hat{\beta}^M_1$	$\hat{\beta}^M_2$	$\hat{c}^M$	CP for $\hat{\beta}^M_1$	CP for $\hat{\beta}^M_2$	CP for $\hat{c}^M$
Setting 1	−0.002(0.036)	0.0002(0.036)	0.0002(0.023)	95.5%	95.5%	95.3%
Setting 2	−0.009(0.061)	0.003(0.060)	−0.001(0.043)	96.0%	96.0%	93.8%
Setting 3	−0.010(0.091)	−0.002(0.089)	−0.001(0.033)	93.7%	93.7%	94.5%
Setting 4	−0.029(0.136)	0.034(0.130)	−0.002(0.056)	98.3%	98.3%	95.0%

For each setting, we conduct 600 simulations. The biases, standard deviations (SD) of $\hat{\beta}^M$ and $\hat{c}^M$, and coverage probabilities (CP) of 95% Wald-type confidence intervals for $\beta^M_{(0)}$ and $c^M_{(0)}$ are reported in Tables 2. In all scenarios, the proposed estimators achieve the smallest biases and standard deviations in Setting 1, where the baseline function and the propensity score are both correctly specified. In Settings 2 and 3, the proposed estimators are nearly unbiased, showing the doubly robust property of the subgroup estimators obtained using the A-learning estimating equations. In Setting 4, where the baseline function and the propensity score are both misspecified, biases and standard deviations of the estimators tend to be larger, however, the biases are still reasonably small. In addition, the coverage probabilities of 95% Wald-type confidence intervals are close to the nominal level for all cases.

To further assess the performance of the proposed maximin OTRs, we compare it with the estimated pooled OTR, $\hat{d}^P(x) = I(x^T\hat{\beta}^P > -\hat{c}^P)$ and the OTR based on random effects models, $\hat{d}^R(x) = I(x^T\hat{\beta}^R > -\hat{c}^R)$. Here, $\hat{\beta}^P$ and $\hat{c}^P$ are obtained based on pooled data by solving a single A-learning estimating equation. To obtain $\hat{\beta}^R$ and $\hat{c}^R$, we first obtain $\hat{\beta}_g$, $\hat{c}_g$ by solving A-learning estimating equations, based on $\{X_{gj}, A_{gj}, Y_{gj}\}_{j=1}^m$. The covariance of $(\hat{\beta}_g^T, \hat{c}_g)^T$ is estimated by the sandwich estimator. Based on these estimators, we calculate $\hat{\beta}^R$ and $\hat{c}^R$ using the R package `mvmeta`. The between-group covariance matrix is estimated by the method of moments. For both scenarios, we consider the following leave-one-group-out cross-validation procedure for evaluation. We first obtain estimators $\hat{\beta}^M$, $\hat{c}^M$, $\hat{\beta}^P$, $\hat{c}^P$, $\hat{\beta}^R$ and $\hat{c}^R$ based on pooled samples of any three groups. Then, we

evaluate the PCD and the VD as defined in Section 3.1 under the obtained maximin OTR and the pooled OTR for the remaining testing group, using Monte Carlo simulations based on the true model for the testing group.

Table 3 and 4 summarize the results of the VD for Scenario 1 and Scenario 2. The results of the PCD are given in Table 21 and 22 in the supplementary article. The OTR obtained by random effects meta-analyses is close to the estimated pooled OTR in both scenarios. In Scenario 1, both the PCD and the VD under our maximin OTR are much higher than those under the other two OTRs for all the testing groups. Taking PCD as an example, on average, the PCD under the maximin OTR is approximately $5 \sim 6\%$ higher than those under the other OTRs. This demonstrates the advantages of the proposed maximin-projection learning when there is relatively large heterogeneity in optimal treatment decision-making across subgroups. In Scenario 2, since the groupwise optimal treatment regimes are "close" to each other in "angles", all the estimated OTRs do not differ much. From Table 4, it can be seen that our maximin OTR performs better than the other OTRs when the first group is taken as the testing group, while it has comparable performance with the other OTRs for other groups as testing groups.

In Section C.2 in the supplementary article, we conduct some additional simulation experiments with non-normal covariates. Findings are similar to those with normal covariates.

Table 3: VD results (with standard errors in parenthesis) for Scenario 1 under the estimated maximin OTR $\hat{d}_M$, the pooled OTR $\hat{d}_P$ and the OTR obtained by random effects meta-analyses $\hat{d}_R$.

Testing group		First group	Second group	Third group	Fourth group
Setting 1	$\hat{d}_P$	0.407(0.002)	0.606(0.001)	0.632(0.002)	0.368(0.002)
	$\hat{d}_R$	0.408(0.001)	0.608(0.001)	0.633(0.001)	0.367(0.001)
	$\hat{d}_M$	0.486(0.001)	0.690(0.001)	0.723(0.001)	0.458(0.001)
Setting 2	$\hat{d}_P$	0.406(0.002)	0.606(0.002)	0.630(0.002)	0.366(0.002)
	$\hat{d}_R$	0.407(0.001)	0.608(0.001)	0.633(0.001)	0.366(0.001)
	$\hat{d}_M$	0.483(0.002)	0.689(0.001)	0.719(0.001)	0.452(0.002)
Setting 3	$\hat{d}_P$	0.407(0.003)	0.604(0.002)	0.630(0.002)	0.367(0.003)
	$\hat{d}_R$	0.405(0.002)	0.606(0.001)	0.632(0.001)	0.367(0.002)
	$\hat{d}_M$	0.483(0.002)	0.688(0.001)	0.723(0.001)	0.454(0.002)
Setting 4	$\hat{d}_P$	0.406(0.003)	0.602(0.003)	0.628(0.003)	0.365(0.003)
	$\hat{d}_R$	0.406(0.002)	0.606(0.001)	0.632(0.001)	0.366(0.002)
	$\hat{d}_M$	0.473(0.003)	0.686(0.002)	0.716(0.001)	0.439(0.004)

Table 4: VD results (with standard errors in parenthesis) for Scenario 2 under the estimated maximin OTR $\hat{d}_M$, the pooled OTR $\hat{d}_P$ and the OTR obtained by random effects meta-analyses $\hat{d}_R$.

Testing group		First group	Second group	Third group	Fourth group
Setting 1	$\hat{d}_P$	0.803(<0.001)	0.597(<0.001)	0.865(<0.001)	0.762(<0.001)
	$\hat{d}_R$	0.803(<0.001)	0.598(<0.001)	0.865(<0.001)	0.761(<0.001)
	$\hat{d}_M$	0.847(<0.001)	0.588(<0.001)	0.865(<0.001)	0.769(<0.001)
Setting 2	$\hat{d}_P$	0.802(0.001)	0.597(<0.001)	0.864(<0.001)	0.761(<0.001)
	$\hat{d}_R$	0.803(<0.001)	0.598(<0.001)	0.865(<0.001)	0.762(<0.001)
	$\hat{d}_M$	0.843(0.001)	0.587(<0.001)	0.863(<0.001)	0.767(0.001)
Setting 3	$\hat{d}_P$	0.801(0.001)	0.597(<0.001)	0.863(<0.001)	0.760(0.001)
	$\hat{d}_R$	0.801(0.001)	0.597(<0.001)	0.864(<0.001)	0.761(0.001)
	$\hat{d}_M$	0.841(0.001)	0.588(<0.001)	0.861(0.001)	0.765(0.001)
Setting 4	$\hat{d}_P$	0.799(0.001)	0.595(<0.001)	0.861(0.001)	0.758(0.001)
	$\hat{d}_R$	0.804(0.001)	0.597(<0.001)	0.863(<0.001)	0.759(0.001)
	$\hat{d}_M$	0.826(0.002)	0.587(0.001)	0.853(0.001)	0.756(0.002)

Although our maximin estimators have better performance for treatment decision making in the above simulation examples, they can have larger variances compared with the random effects models. This is a potential disadvantage of our method.

6. Health assessment questionnaire progression data

The health assessment questionnaire progression data comes from an observational study to investigate the influence of early disease modifying antirheumatic drug (DMARD) treatment and its duration for patients with recent onset inflammatory polyarthritis (Farragher et al., 2010). Early DMARDs treatment was routinely used in the management of rheumatoid arthritis (RA). Among conventional DMARDs, Methotrexate is the most widely used one and is now considered a benchmark against new treatments to be used. Previous studies showed that RA patients who have failed to respond to methotrexate may have clinically important improvements if treated with combination DMARDs, such as methotrexate-sulfasalazine-hydroxychloroquine, methotrexate-sulfasalazine-steroids or other Methotrexate combinations (Boers et al., 1997). However, Methotrexate combinations did not work for all RA patients, and they may not add benefits in some patients who were stable on DMARD monotherapy (Symmons et al., 2005). It is of clinical interest to develop individualized OTRs and to know which patients will benefit from treating with Methotrexate combinations. The study sample include 420 patients who were recruited to the study from 1990 to 2000 and were treated with either

methotrexate monotherapy or methotrexate combinations. Age, gender, duration of disease, HAQ score, number of swollen joints and number of tender joints were recorded at baseline. We standardize all six covariates such that their sample covariance matrix equals the identity matrix within each group. We compare methotrexate combinations ($A = 1$) with methotrexate monotherapy ($A = 0$). The difference HAQ scores between baseline and 5-year is set to be the response. Here, we classify 420 patients into three groups according to their recruitment time. Specifically, group 1 includes patients enrolled from 1990 to 1992; group 2 includes those enrolled from 1993 to 1996; and group 3 includes those enrolled from 1997 to 2000. Sample sizes of the three groups are 265, 78 and 77, respectively.

In our analysis, we use the last two standardized covariates to fit the contrast function, since the regression coefficients of other variables are not significant. Denoted these two covariates by $X_{gj}^{(1)}$ and $X_{gj}^{(2)}$, respectively. For each group g, we consider the following model

$$\mathrm{E}(Y_{gj}|X_{gj}, A_{gj}) = h_g(X_{gj}) + A_{gj}(c_0 + \beta_{g1}X_{gj}^{(1)} + \beta_{g2}X_{gj}^{(2)}).$$

The parameters $c_0, \beta_{g1}, \beta_{g2}$ are estimated using the A-learning estimating equations as discussed in Section 4.2. Here, a linear model is fitted for the baseline function and a logistic regression model is fitted for the propensity score. When fitting the propensity score model, all six covariates are included. Table 5 reports the group-wise estimators obtained using the A-learning estimating equations, suggesting there is some heterogeneity in optimal treatment regimens across three groups.

Table 5: Estimators of groupwise OTR (standard errors in paranthesis) for the HAQ data.

	Group 1	Group 2	Group 3
$\hat{\beta}_{g1}$	0.05(0.11)	−0.40(0.17)	0.07(0.21)
$\hat{\beta}_{g2}$	0.07(0.11)	0.06(0.21)	0.32(0.16)

Table 6: $\hat{d}_M$, $\hat{d}_P$, $\hat{d}_R$ and their value functions

Testing group	Group 1			Group 2			Group 3		
	$\hat{d}_M$	$\hat{d}_P$	$\hat{d}_R$	$\hat{d}_M$	$\hat{d}_P$	$\hat{d}_R$	$\hat{d}_M$	$\hat{d}_P$	$\hat{d}_R$
$\hat{c}$	−0.87	−0.14	−0.12	−2.38	−0.21	−0.11	−3.08	−0.31	−0.32
$\hat{\beta}_1$	−0.48	−0.02	−0.00	0.61	0.16	0.16	−0.02	0.06	−0.01
$\hat{\beta}_2$	0.88	0.25	0.23	0.79	0.10	0.14	1.00	0.06	0.10
$\hat{\mathrm{E}}Y_g^\star(d)$	−0.08	−0.09	−0.09	−0.09	−0.19	−0.22	−0.12	−0.13	−0.12

We use the same leave-one-group-out cross validation procedure as done in simulations to evaluate the performance of the proposed method. We calculate the maximin OTR $\hat{d}_M$, the pooled OTR $\hat{d}_P$, and the OTR obtained by random effects meta-analyses $\hat{d}_R$ based on every two groups of patients, and

evaluate them on the remaining group based on the estimated value function. For a given treatment regime d and group g, the estimated value function is given by

$$\hat{\mathrm{E}}Y_g^\star(d) = \frac{1}{m_g}\sum_{j=1}^{m_g}\left[Y_{gj} + \left(\hat{c}_0 + \hat{\beta}_{g1}X_{gj}^{(1)} + \hat{\beta}_{g2}X_{gj}^{(2)}\right)\{d(X_{gj}) - A_{gj}\}\right],$$

which is computed based on the advantage function as introduced in Murphy (2003). Results are given in Table 6. Value under the maximin OTR are uniformly better than those under other OTRs across all three groups, showing a big improvement for group 2. Besides, the estimators involved in the regimes $\hat{d}_P$ and $\hat{d}_R$ are very close.

7. Discussion

In this paper, we propose a maximin-projection learning to aggregate OTRs for patients from different populations with heterogeneity. It has appealing statistical interpretations in the sense of maximizing the minimum PCD and the minimum value difference across subgroups. The corresponding estimation procedure is easy to implement via quadratically constrained linear programming, and the asymptotic properties of the resulting estimators are studied.

7.1. Alternative maximin formulation

Our procedure requires to scale the baseline covariates X_g to mean zero and identity covariance matrix for $g = 1, 2, \ldots, G, G+1$. Let $X_{g,0}$ be the original variable prior to transformation and $\beta_{g,0}$, $c_{g,0}$ the corresponding individualized and marginal treatment effects, respectively. The proposed maximin OTR is constructed based on $\beta^M = \arg\max_{\|\beta\|_2=1}\min_{g\in\{1,\ldots,G\}}\beta^T\beta_g$, or equivalently,

$$\beta^{M*} = \arg\max_{\|\Sigma_{G+1}^{1/2}\beta\|_2=1}\min_{g\in\{1,\ldots,G\}}\beta^T\Sigma_{G+1}^{1/2}\Sigma_g^{1/2}\beta_{g,0},$$

where Σ_g is the covariance matrix of $X_{g,0}$ for $g = 1, \ldots, G+1$.

As pointed by one of the referee, we can also consider the maximin OTR based on β^{M**} where

$$\beta^{M**} = \arg\max_{\|\Sigma_{G+1}^{1/2}\beta\|_2=1}\min_{g\in\{1,\ldots,G\}}\beta^T\Sigma_{G+1}\beta_{g,0}.$$

Assuming $\mathrm{E}X_{1,0} = \mathrm{E}X_{2,0} = \cdots = \mathrm{E}X_{G,0} = \mathrm{E}X_{G+1,0} = 0$, $c_{1,0} = c_{2,0} = \cdots = c_{G,0}$ and X_{G+1} is spherically distributed, we can show

$$\beta^{M**} = \arg\max_{\|\Sigma_{G+1}^{1/2}\beta\|_2=1}\min_{g\in\{1,\ldots,G\}}\mathrm{E}(X_{G+1,0}^T\beta_{g,0} + c_{g,0})I(X_{G+1,0}^T\beta + c),$$

for any $c > 0$. This implies that β^{M**} maximizes the minimum groupwise value difference function under the new distribution $X_{G+1,0}$.

It is worthwhile to investigate the performance of the OTR based on β^{M**}. However, this is beyond the scope of the current paper. Below, we briefly compare the proposed maximin OTR with the maximin OTR based on β^{M**} and discuss their connections. First, β^{M**} maximizes the minimum groupwise value difference function under the new distribution $X_{G+1,0}$ while β^M maximizes the minimum groupwise value difference function under the new distribution X_{G+1} after scaling. To see this, note that when $c_1 = \cdots = c_G$ and X_{G+1} is spherically distributed, we have

$$\beta^M = \arg \max_{\|\beta\|_2=1} \min_{g\in\{1,\ldots,G\}} \mathrm{E}(X_{G+1}^T\beta_g + c_g)I(X_{G+1}^T\beta + c),$$

for any $c > 0$. Second, β^{M**} usually doesn't coincide with β^{M*}. A sufficient condition for $\beta^{M*} = \beta^{M**}$ is that $\Sigma_1 = \Sigma_2 = \cdots = \Sigma_G = \Sigma_{G+1}$. Lastly, estimating β^{M**} might exhibit less variances than β^{M*}, since it doesn't require the estimation of $\Sigma_1, \ldots, \Sigma_G$. However, the OTR based on β^{M**} is not scale invariant. To see this, let $X_{G+1}^{**} = CX_{G+1,0}$ for some invertible matrix C. The covariance matrix of $X_{G+1,0}$ is equal to $C\Sigma_{G+1}C^T$. Let

$$\beta^{M***} = \arg \max_{\|\Sigma_{G+1}^{1/2}C^T\beta\|_2=1} \min_{g\in 1,\ldots,G} \beta^T C\Sigma_{G+1}C^T\beta_{g,0},$$

there's no guarantee that $\beta^{M***} = (C^T)^{-1}\beta^{M**}$.

7.2. Extensions

In current work, we mainly deal with heterogeneity caused by groupwise individualized treatment effects β_g's, and assume the same marginal treatment effects c_g for all groups. It is possible to extend our proposed maximin projection learning to the case when c_g's vary across different groups as well. Specifically, consider

$$(\hat{\beta}^M, \hat{c}^M) = \arg \max_{\|\beta\|_2^2+c^2\le 1} \min_{g\in\{1,\ldots,G\}} \left(\hat{\beta}_g^T\beta + \hat{c}_g c\right),$$

where $\hat{\beta}_g$ and $\hat{c}_g$ are subgroup estimators. Statistical properties of $\hat{\beta}^M$ and $\hat{c}^M$ can be similarly established. For example, $\hat{\beta}^M$ and $\hat{c}^M$ can be shown to converge almost surely to some $\beta_{(0)}^M$ and $c_{(0)}^M$, respectively. However, the defined $\beta_{(0)}^M$ and $c_{(0)}^M$ can no longer preserve the interpretation of maximizing the minimum PCD and the minimum VD, due to the fact that the PCD and the VD are complicated functions of (β_g, c_g) and (β, c) when c_gs vary across groups. Consequently, the angle interpretation as demonstrated by the toy example given in Section 2.2 does not hold.

To establish the consistency and asymptotic normality of $\hat{\beta}^M$ and $\hat{c}^M$, we require $\beta_g, g \in K_0$ to be linearly independent. In Section C.1 in the supplementary article, we conduct some additional simulation studies to examine our methods under settings where some of the β_g's are the same. Results suggest that $\hat{\beta}^M$ and $\hat{c}^M$ are still consistent to $\beta_{(0)}^M$ and $c_{(0)}^M$, in these settings. We further

evaluate the VD and the PCD under the estimated maximin OTR and compare them with those under the estimated pooled OTR. Findings are similar to those in Section 5.

In addition, in our current work, we assume a linear interaction between treatment and baseline covariates. It is interesting to consider a more general model as follows:

$$Y_g = h_g(X_g) + A_g Q(\beta_g^T X_g + c_g) + e_g, \qquad g = 1, \ldots, G, \tag{14}$$

where Q is a strictly monotone increasing function with $Q(0) = 0$. The parameters β_g in each group can be consistently estimated using the concordance-assisted learning method by Fan et al. (2017). The properties of the corresponding maximin-projection estimator warrant further investigation.

Acknowledgements

We thank the editor, the AE and three referees for providing helpful suggestions that significantly improved the quality of the paper. The research of Chengchun Shi and Rui Song is partially supported by Grant NSF-DMS-1555244 and Grant NCI P01 CA142538. The research of Wenbin Lu is partially supported by Grant NCI P01 CA142538.

References

Avi-Itzhak, H., J. A. Van Mieghem, and L. Rub (1995). Multiple subclass pattern recognition: a maximin correlation approach. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 17*(4), 418–431.

Boers, M., A. C. Verhoeven, H. M. Markusse, M. A. van de Laar, R. Westhovens, J. C. van Denderen, D. van Zeben, B. A. Dijkmans, A. J. Peeters, P. Jacobs, et al. (1997). Randomised comparison of combined step-down prednisolone, methotrexate and sulphasalazine with sulphasalazine alone in early rheumatoid arthritis. *The Lancet 350*(9074), 309–318.

Bühlmann, P. and N. Meinshausen (2016). Magging: maximin aggregation for inhomogeneous large-scale data. *Proceedings of the IEEE 104*(1), 126–135.

Chakraborty, B., S. Murphy, and V. Strecher (2010). Inference for non-regular parameters in optimal dynamic treatment regimes. *Stat. Methods Med. Res. 19*(3), 317–343.

Chen, H., A. K. Manning, and J. Dupuis (2012). A method of moments estimator for random effect multivariate meta-analysis. *Biometrics 68*(4), 1278–1284.

DerSimonian, R. and N. Laird (1986). Meta-analysis in clinical trials. *Controlled clinical trials 7*(3), 177–188.

Dunn, G. and R. Bentall (2007). Modelling treatment-effect heterogeneity in randomized controlled trials of complex interventions (psychological treatments). *Statistics in medicine 26*(26), 4719–4745.

Fan, C., W. Lu, R. Song, and Y. Zhou (2017). Concordance-assisted learning for estimating optimal individualized treatment regimes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 79*(5), 1565–1582.

Fang, K. T., S. Kotz, and K. W. Ng (1990). *Symmetric multivariate and related distributions*, Volume 36 of *Monographs on Statistics and Applied Probability*. Chapman and Hall, Ltd., London.

Farragher, T. M., M. Lunt, B. Fu, D. Bunn, and D. P. Symmons (2010). Early treatment with, and time receiving, first disease-modifying antirheumatic drug predicts long-term function in patients with inflammatory polyarthritis. *Annals of the rheumatic diseases 69*(4), 689–695.

Jackson, D., I. R. White, and S. G. Thompson (2010). Extending DerSimonian and Laird's methodology to perform multivariate random effects meta-analyses. *Stat. Med. 29*(12), 1282–1297.

Lee, T., T. Moon, S. J. Kim, and S. Yoon (2016). Regularization and kernelization of the maximin correlation approach. *IEEE Access 4*, 1385–1392.

Meinshausen, N. and P. Bühlmann (2015). Maximin effects in inhomogeneous large-scale data. *Ann. Statist. 43*(4), 1801–1830.

Murphy, S. A. (2003). Optimal dynamic treatment regimes. *J. R. Stat. Soc. Ser. B Stat. Methodol. 65*(2), 331–366.

Robins, J., M. Hernan, and B. Brumback (2000). Marginal structural models and causal inference in epidemiology. *Epidemiol. 11*, 550–560.

Rubin, D. (1974). Estimating causal effects of treatments in randomized and non-randomized studies. *J. Edu. Psychol. 66*, 688–701.

Symmons, D., K. Tricker, C. Roberts, L. Davies, P. Dawes, and D. Scott (2005). The british rheumatoid outcome study group (brosg) randomised controlled trial to compare the effectiveness and cost-effectiveness of aggressive versus symptomatic therapy in established rheumatoid arthritis. *Health technology assessment (Winchester, England) 9*(34), iii–iv.

Tarrier, N., S. Lewis, G. Haddock, R. Bentall, R. Drake, P. Kinderman, D. Kingdon, R. Siddle, J. Everitt, K. Leadley, et al. (2004). Cognitive-behavioural therapy in first-episode and early schizophrenia. *The British Journal of Psychiatry 184*(3), 231–239.

Watkins, C. and P. Dayan (1992). Q-learning. *Mach. Learn. 8*, 279–292.

Zhang, B., A. A. Tsiatis, E. B. Laber, and M. Davidian (2012). A robust method for estimating optimal treatment regimes. *Biometrics 68*(4), 1010–1018.

Zhao, Y., D. Zeng, A. J. Rush, and M. R. Kosorok (2012). Estimating individualized treatment rules using outcome weighted learning. *J. Amer. Statist. Assoc. 107*(499), 1106–1118.

A. Proof of theorem 2

Before proving theorem 2, we state the following lemma whose proof is given in Section F of the on-line supplementary article.

LEMMA 4. *Consider a set of vectors $\beta_1, \ldots, \beta_G$ of dimension s and function $h(\beta, b, t)$ defined on the domain $\left\{(\beta^T, b^T, t) \in \mathbb{R}^s \times \mathbb{R}^s \times \mathcal{T} : \|\beta\|_2 = 1\right\}$. Assume $h(\beta, b, t) = g(\beta^T b, t)$ for some function $g(\cdot, \cdot)$. Besides, assume for any fixed t, $g(c, t)$ is monotonically increasing as a function of c. Then, for any random variable T defined on $\mathcal{T}$, we have*

$$\arg \max_{\substack{\beta \in \mathbb{R}^p \\ \|\beta\|_2 = 1}} \min_g Eh(\beta_g, \beta, T) = \arg \max_{\substack{\beta \in \mathbb{R}^p \\ \|\beta\|_2 = 1}} \min_g {\beta_g}^T \beta.$$

Since X_g's are identically distributed for $g = 1, \ldots, G$, we omit the subscript g for brevity. We need to show β^M maximizes

$$\min_g \left[\mathrm{E}\left\{h_g(X) + (X^T\beta_g + c_0)I(X^T\beta > -c)\right\} - \mathrm{E}h_g(X)\right] = \min_g \mathrm{E}(X^T\beta_g + c_0)I(X^T\beta > -c).$$

We first show for any $\|\beta\|_2 = 1$ and c, the probability $\Pr(X^T\beta > c)$ is constant as a function of β. Since $\|\beta_g\|_2 = 1$, it follows from lemma F.1 that there exists some orthogonal matrix U such that $U\beta = e_1 = (1, 0, \ldots, 0)^T$. Hence

$$\Pr(X^T\beta > c) = \Pr(X^T U^T U\beta > c) = \Pr(X^T U^T e_1 > c) = \Pr(X^T e_1 > c), \tag{15}$$

where the last equality is due to the definition of spherical distribution (see Definition F.1). By (15), it suffices to show β^M maximizes

$$\min_g \mathrm{E}(X^T\beta_g)I(X^T\beta > -c). \tag{16}$$

Let $\rho_g = \beta^T\beta_g/\|\beta_g\|_2$. Since X is spherically distributed, we have

$$\mathrm{E}(X^T\beta_g)I(X^T\beta > -c) = \|\beta_g\|_2 \mathrm{E}\left(\rho_g X^{(1)} + \sqrt{1-\rho_g^2}X^{(2)}\right) I\left(X^{(1)} > -c\right), \tag{17}$$

for all β_g, c and β such that $\|\beta\|_2 = 1$, where $X^{(1)}$ and $X^{(2)}$ are the first two components of the random vector X. It follows from theorem 2.6 in Fang et al. (1990) that

$$(X^{(1)}, X^{(2)}) \stackrel{d}{=} rd(U_1, U_2), \tag{18}$$

with $r = \|X\|_2$, $d \sim B(1, p/2-1)$, U_1 and U_2 uniformly distributed on the surface $u_1^2 + u_2^2 = 1$, where $B(p,q)$ stands for the Beta distribution with parameters p, q. The random variables r, d are independent of U_1 and U_2. Set $T = rd$. Combining (17) with (18) gives

$$\begin{aligned} & \mathrm{E}(X^T\beta_g)I(X^T\beta > -c) = \mathrm{E}\|\beta_g\|_2 T(\rho_g U_1 + \sqrt{1-\rho_g^2}U_2)I(TU_1 > -c) \qquad (19) \\ = \;& \mathrm{E}\left[\left\{\mathrm{E}\|\beta_g\|_2 t(\rho_g U_1 + \sqrt{1-\rho_g^2}U_2)I(U_1 > -c/t)\right\}|T=t\right] \equiv \mathrm{E}[h(\beta,\beta_g,t)|T=t], \end{aligned}$$

for any β and c such that $\|\beta\|_2 = 1$.

When $c/t \le -1$, we have $I(U_1 > -c/t) = I(U_1 > 1) = 0$ and hence $h = 0$. When $c/t \ge 1$,

$$h(\beta,\beta_g,t) = t\|\beta_g\|_2 \mathrm{E}(\rho_g U_1 + \sqrt{1-\rho_g^2}U_2) = 0.$$

Obviously, in these two trivial cases, h is an increasing function of $\beta^T\beta_g$. Now we consider the case where $c/t = \cos(\psi_1)$ for some $\psi_1 \in (0,\pi)$. Assume $\rho_g = \cos(\psi_2)$ for some $\psi_2 \in (0,\pi)$. The function h can further be simplified to

$$\begin{aligned} & h(\beta,\beta_g,t) = \frac{1}{2\pi}\int_{-\psi_1}^{\psi_1} \|\beta_g\|_2 t\{\cos(\psi_2)\cos(\psi) + \sin(\psi_2)\sin(\psi)\}d\psi \\ = \;& \frac{1}{2\pi}t\|\beta_g\|_2 \int_{-\psi_1}^{\psi_1} \cos(\psi-\psi_2)d\psi = \frac{1}{\pi}t\|\beta_g\|_2 \sin(\psi_1)\cos(\psi_2) = \frac{1}{\pi}t\sin(\psi_1)\beta^T\beta_g. \end{aligned}$$

This proves h is an increasing function of $\beta^T\beta_g$. Hence, (16) follows by an application of lemma 4.